**Danh sách đề tài (N04\_K62)**

|  |  |  |  |
| --- | --- | --- | --- |
| **STT** | **Đề tài** | **Tài liệu** | **Nhóm** |
| 1 | Khai phá luật kết hợp  (Giải thuật Apriori) | Chapter 6 | 13,12,11 |
| 2 | Khai phá luật kết hợp  (Giải thuật FP-Growth) | Chapter 6 | 19,16,1 |
| 3 | Phân loại dữ liệu với cây quyết định (Decision tree – ID3) | Chapter 8 | 15,14,2 |
| 4 | Phân loại dữ liệu với cây quyết định  (Decision tree – C4.5) | Chapter 8 | 9,4 |
| 5 | Phân loại dữ liệu với mạng Bayesian (Naïve Bayesian) | Chapter | 10,7,5 |
| 6 | Phân cụm dữ liệu: K-Means | Chapter 10 | 20,18,3 |
| 7 | Phân cụm dữ liệu: K-Medoid | Chapter 10 | 17,8,6 |

**Công cụ**

**WEKA:** <http://www.cs.waikato.ac.nz/ml/weka/>

**KNIME:** <https://www.knime.com/>

**R:** [https://www.r-project.org](https://www.r-project.org/)

**Python**

**Tập dữ liệu**

1. UCI Machine Learning Repository

<https://archive.ics.uci.edu/datasets>

1. Kaggle

<https://www.kaggle.com/datasets>

3. Microsoft <https://msropendata.com/datasets?domain=COMPUTER%20SCIENCE>

**BÁO CÁO ĐỀ TÀI KHAI PHÁ DỮ LIỆU**

1. Tổng quan về Khai phá luật kết hợp | Phân lớp dữ liệu | Phân cụm dữ liệu

2. Thuật toán :

+ Mô tả thuật toán

+ Ví dụ :

3. Ứng dụng

+ Tên ứng dụng

+ Nguồn dữ liệu

+ Cách thức thực hiện

+ Kết quả

***Yêu cầu trình bày báo cáo***

Font chữ : Time New Roman, Font size : 13

Paper size: A4, Left: 3 cm, Right: 2, Top: 2 cm, Bottom: 2cm

Line Spacing: 1.2 Before: 6pt, Apter: 6pt

Số trang : 10 - 15 trang

Tên file: **KPDL\_N04\_NhomXX\_Baocao.docx**

**Deadline: 16/11/2024 (Báo cáo, Dữ liệu, Code (nếu có))**

**MS Teams => Đề tài => NHOMXX**

Lập kế hoạch: KPDL\_N04\_NhomXX.docx

- Báo cáo lý thuyết vào thứ 5 ngày 17/10/2024

- Nộp slide báo cáo: 16/10/2024

- Báo cáo ứng dụng: từ ngày 21/10/2024